Published for SISSA by Springer

Received: *January 27, 2018*
Accepted: *April 8, 2018*
Published: *April 11, 2018*

Phase diagram of q-deformed Yang-Mills theory on S^2 at non-zero θ-angle

Kazumi Okuyama

Department of Physics, Shinshu University,
Matsumoto 390-8621, Japan

E-mail: kazumi@azusa.shinshu-u.ac.jp

Abstract: We study the phase diagram of q-deformed Yang-Mills theory on S^2 at non-zero θ-angle using the exact partition function at finite N. By evaluating the exact partition function numerically, we find evidence for the existence of a series of phase transitions at non-zero θ-angle as conjectured in [hep-th/0509004].

Keywords: 1/N Expansion, Matrix Models, Topological Strings

ArXiv ePrint: 1801.08236

https://doi.org/10.1007/JHEP04(2018)059

Contents

1 Introduction

The topological θ-angle plays an important role in the dynamics of gauge theories and σ-models in various dimensions (see e.g. [1–5] and references therein). More than a decade ago, the phase diagram of q-deformed Yang-Mills theory (qYM) on S^2 at non-zero θ-angle was conjectured in [6]. In this paper, we will study the phase diagram of qYM on S^2 using the exact partition function at finite N. In general, exact result is a very powerful tool to analyze the non-perturbative aspects of gauge theories, which are usually beyond reach by other means. This approach was successfully applied to many examples, including the ABJM theory on S^3 [7–10] and the Gross-Witten-Wadia (GWW) unitary matrix model [11–13].

In [6, 14, 15], it was found that qYM on S^2 in the large N 't Hooft limit has a third order phase transition at some critical value $t = t_c$ of the 't Hooft coupling when $\theta = 0$. The physical mechanism of this phase transition is essentially the same as that of ordinary (undeformed) Yang-Mills theory on S^2 found by Douglas and Kazakov [16]. As shown in [6, 14, 15], at the critical point the instanton contribution becomes comparable with the zero-instanton sector and the phase transition is triggered by the instanton condensation, in much the same way as the undeformed Yang-Mills theory [17].

In [6], it was further conjectured that the phase diagram of qYM at non-zero θ has an intricate structure (see figure 3 in [6]): there is a series of phase transition curves on the t-θ plane which accumulate at the point $(t, \theta) = (0, \pi)$. Each transition curve corresponds to the exchange of dominance of different instanton sectors. In this paper we will examine this conjecture using the exact partition function of qYM at finite N. By evaluating the exact partition function numerically, we find evidence for the conjectured phase diagram in [6]. We find that the *prefactor* of instanton contribution is important for the understanding of the phase structure at non-zero θ. However, it turns out that it is difficult to find the analytic form of this prefactor in the q-deformed case. Instead, in section 5 we consider the undeformed case where the prefactor of 1-instanton contribution is already known [17], and we propose an analytic form of the first two transition curves at non-zero θ and check this proposal numerically.

This paper is organized as follows. In section 2, we write down the exact partition function of qYM on S^2 as a determinant of $N\times N$ matrix. Using this exact result, we study the behavior of free energy at $\theta=0$ and confirm the phase transition found in [6, 14, 15]. In section 3, we study the instanton contribution using our exact result at finite N. In section 4, we numerically compute the free energy at non-zero θ and find evidence that indeed there is a series of phase transitions coming from the exchange of dominance of different instanton sectors, as conjectured in [6]. In section 5, we consider the phase diagram of undeformed theory at $\theta\neq 0$ and propose an analytic form of the first two transition curves. Finally, we conclude in section 6 with some discussion of the future directions.

2 Exact partition function at finite N

In this section, we consider the exact partition function of qYM on S^2 at finite N. This section is mostly a review of the known results.

The q-deformed U(N) Yang-Mills theory on S^2 naturally appears as a worldvolume theory on N D4-branes on a non-compact Calabi-Yau X_p

$$X_p:\ \mathcal{O}(-p)\oplus\mathcal{O}(p-2)\to\mathbb{P}^1, \tag{2.1}$$

where the D4-branes in question are wrapping around the base $\mathbb{P}^1$ and one of the fiber $\mathcal{O}(-p)$ of X_p [18]. It is argued in [18, 19] that the path integral of D4-brane worldvolume theory localizes to the q-deformed Yang-Mills theory on the base $\mathbb{P}^1=S^2$. This partition function of D4-brane theory is identified as the partition function of 4d black holes made of the bound states of D4, D2, D0-branes, which in turn is related to the topological string partition function on X_p via the OSV conjecture [20].

It is known that qYM on any Riemann surface can be solved exactly [18] in a similar manner as the ordinary undeformed 2d Yang-Mills theory [21]. The partition function of U(N) qYM on S^2 is given by

$$Z_N=\frac{1}{N!}\sum_{n_i\in\mathbb{Z}+\frac{\epsilon}{2}}\prod_{i<j}[n_i-n_j]^2\exp\left(-\frac{g_s p}{2}\sum_{i=1}^N n_i^2+\mathrm{i}\theta\sum_{i=1}^N n_i\right), \tag{2.2}$$

where $[n]$ denotes the q-integer

$$[n]=\mathbf{q}^{\frac{n}{2}}-\mathbf{q}^{-\frac{n}{2}},\qquad \mathbf{q}=e^{-g_s}, \tag{2.3}$$

and ϵ is given by

$$\epsilon = \begin{cases} 0, & (\text{odd } N), \\ 1, & (\text{even } N). \end{cases} \tag{2.4}$$

In other words, the summation of n_i in (2.2) runs over integers for odd N and half-integers for even N.[1] The overall normalization of the partition function is ambiguous; in (2.2) we followed the convention of [6]. As we will show below, we can rewrite this partition function as a determinant of $N \times N$ matrix. To see this, we first notice that the factor $\prod_{i<j}[n_i - n_j]$ in (2.2) is basically the Vandermonde determinant and it is rewritten as

$$\prod_{i<j}[n_i - n_j] = \sum_{\sigma \in S_N} (-1)^\sigma \prod_{i=1}^{N} \mathbf{q}^{(i-\frac{N+1}{2})n_{\sigma(i)}}. \tag{2.6}$$

This relation can also be understood as the Weyl denominator formula of $\mathrm{U}(N)$ gauge group. Squaring the above expression (2.6) we get a sum over two permutations, but one of them can be trivialized by using the invariance of $\sum_i n_i^2$ and $\sum_i n_i$ under the permutation of n_i. In this way (2.2) is rewritten as

$$\begin{aligned} Z_N &= \sum_{n_i \in \mathbb{Z}+\frac{\epsilon}{2}} \sum_{\sigma \in S_N} (-1)^\sigma \prod_{i=1}^{N} \mathbf{q}^{(i+\sigma(i)-N-1)n_i} e^{-\frac{1}{2}g_s p n_i^2 + \mathrm{i}\theta n_i} \\ &= \sum_{\sigma \in S_N} (-1)^\sigma \prod_{i=1}^{N} \vartheta_{3-\epsilon}\left(\frac{\theta + \mathrm{i}g_s(i+\sigma(i)-N-1)}{2\pi}, \frac{\mathrm{i}g_s p}{2\pi}\right), \end{aligned} \tag{2.7}$$

where the Jacobi theta function is defined by

$$\vartheta_{3-\epsilon}(v,\tau) = \sum_{n \in \mathbb{Z}+\frac{\epsilon}{2}} e^{2\pi \mathrm{i} n v + \pi \mathrm{i} n^2 \tau}. \tag{2.8}$$

Finally, the sum over S_N in (2.7) reduces to the determinant

$$Z_N = \det\left[\vartheta_{3-\epsilon}\left(\frac{\theta + \mathrm{i}g_s(i+j-N-1)}{2\pi}, \frac{\mathrm{i}g_s p}{2\pi}\right)\right]_{i,j=1,\cdots,N}. \tag{2.9}$$

This is the main result of this section.[2] This determinant form of Z_N is reminiscent of the exact partition function of GWW model [22, 23]. However, there is an important difference: the exact partition function of qYM in (2.9) is given by a Hankel determinant $\det(a_{i+j})$, while the exact partition function of GWW model is given by a Toeplitz determinant $\det(b_{i-j})$.

We are interested in the behavior of this partition function in the large N 't Hooft limit

$$g_s \to 0, \quad N \to \infty, \quad t = g_s N : \text{fixed}, \tag{2.10}$$

[1] Our n_i is related to the rows $\{\lambda_1, \cdots, \lambda_N\}$ of Young diagram by

$$n_i = \lambda_i - i + \frac{N+1}{2}. \tag{2.5}$$

It follows that $n_i \in \mathbb{Z} + \frac{\epsilon}{2}$ since λ_i is integer.

[2] As far as we know, this expression has not appeared in the literature before.

JHEP04(2018)059

and we would like to study the genus expansion of free energy

$$F = \log Z_N = \sum_{g=0}^{\infty} g_s^{2g-2} F_g(t). \tag{2.11}$$

In [6, 14, 15], it is found that when $\theta = 0$ there is a third order phase transition at the critical value $t = t_c$ of the 't Hooft coupling, where t_c is given by

$$t_c = -2p \log\left(\cos\frac{\pi}{p}\right). \tag{2.12}$$

This phase transition occurs only for $p > 2$ [6, 14, 15], and we will assume $p > 2$ throughout this paper. In the rest of this section, we will consider the behavior of free energy above ($t > t_c$) and below ($t < t_c$) the phase transition using the exact partition function Z_N at finite N.

2.1 Strong coupling phase

Let us first consider the strong coupling phase ($t > t_c$). In the large N limit, the eigenvalue distribution in this phase is described by a two-cut solution of a certain matrix model and the explicit form of the resolvent was constructed in [6, 14, 15]. However, it is not so straightforward to compute the genus-zero free energy $F_0(t)$ from this solution.

To study this phase, it is convenient to regard the partition function (2.2) as a sum over configurations of N non-relativistic fermions

$$Z_N = \sum_{n_1 < \cdots < n_N} Z_{\vec{n}} = \sum_{n_1 < \cdots < n_N} \prod_{i<j} [n_i - n_j]^2 e^{-g_s pE + \mathrm{i}\theta P} \tag{2.13}$$

where $\vec{n} = (n_1, \cdots, n_N)$ specifies the momentum of N fermions and the total energy E and the total momentum P of fermions are given by

$$E = \sum_{i=1}^{N} \frac{1}{2} n_i^2, \qquad P = \sum_{i=1}^{N} n_i. \tag{2.14}$$

These N fermions are interacting through the factor $\prod_{i<j}[n_i - n_j]^2$.

In the strong coupling phase, we can compute (2.13) by summing over the fermion configurations from the small total energy E. A similar computation has been performed for the undeformed Yang-Mills case in [16]. One can easily see that the ground state (lowest energy configuration) is given by

$$\vec{n}_0 = (-n_F, -n_F + 1, \cdots, n_F - 1, n_F), \qquad n_F = \frac{N-1}{2}. \tag{2.15}$$

The energy E_0 and the momentum P_0 of the ground state are given by

$$E_0 = \frac{1}{2}\vec{n}_0^2 = \frac{N^3 - N}{24}, \qquad P_0 = 0, \tag{2.16}$$

and the contribution of this ground state is

$$Z_{\text{gnd}} = Z_{\vec{n}_0} = e^{-g_s pE_0} \prod_{i<j} [i-j]^2 = e^{-\frac{g_s p}{24}(N^3 - N)} \prod_{j=1}^{N-1} \left(2 \sinh \frac{g_s j}{2}\right)^{2(N-j)}. \tag{2.17}$$

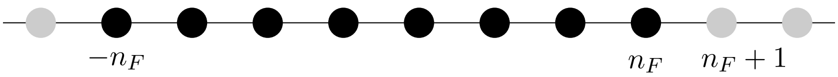


Figure 1. Maya diagram for the ground state. The black nodes ($-n_F \le n \le n_F$) are occupied by fermions while the gray nodes ($|n| > n_F$) are empty.

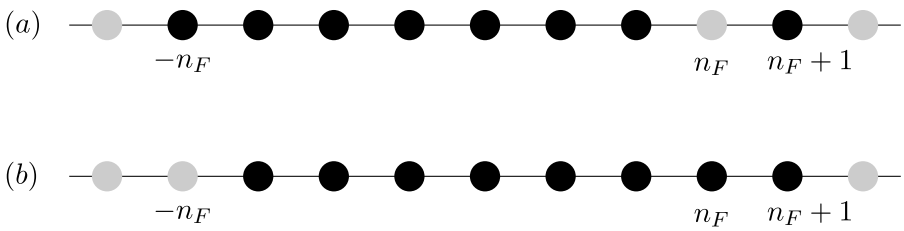


Figure 2. Examples of excited states with $\Delta E = N/2$: (a) chiral, (b) non-chiral.

One can visualize the configuration of fermions by the so-called Maya diagram as shown in figure 1: the black nodes are occupied by the N fermions and the gray nodes are empty. In the ground state the nodes between $n = -n_F$ and $n = n_F$ are occupied; the nodes at $n = \pm n_F$ can be thought of as the Fermi levels.

We can also draw the Maya diagram for excited states as in figure 2. The excitation energy $\Delta E = E - E_0$ and the total momentum of the states (a) and (b) in figure 2 can be easily computed as

$$\begin{aligned}
(a) : \Delta E &= \frac{1}{2}(n_F+1)^2 - \frac{1}{2}n_F^2 = \frac{N}{2}, \qquad && P = n_F + 1 - n_F = 1, \\
(b) : \Delta E &= \frac{1}{2}(n_F+1)^2 - \frac{1}{2}(-n_F)^2 = \frac{N}{2}, \qquad && P = n_F + 1 - (-n_F) = N,
\end{aligned} \tag{2.18}$$

and their contributions to the partition function are given by

$$\frac{Z_{(a)}}{Z_{\text{gnd}}} = \frac{[N]^2}{[1]^2} e^{-\frac{N}{2}g_s p + \mathrm{i}\theta}, \qquad \frac{Z_{(b)}}{Z_{\text{gnd}}} = e^{-\frac{N}{2}g_s p + \mathrm{i}N\theta}. \tag{2.19}$$

There are two more states with the same excitation energy $\Delta E = N/2$ which are obtained by flipping the sign of momenta $n_i \to -n_i$ in figure 2. In this way, we can compute the large t expansion of partition function systematically as

$$\begin{aligned}
\frac{Z_N}{Z_{\rm gnd}} &= 1+\left[2\cos\theta\frac{[N]^2}{[1]^2}+2\cos N\theta\right]e^{-\frac{tp}{2}}\\
&+\left[2\cos 2\theta\frac{\mathbf{q}^p[N]^2[N+1]^2+\mathbf{q}^{-p}[N]^2[N-1]^2}{[1]^2[2]^2}+\frac{[N-1]^2[N+1]^2}{[1]^4}\right.\\
&\left.+\left(2\mathbf{q}^p\cos(N+1)\theta+2\mathbf{q}^{-p}\cos(N-1)\theta\right)\frac{[N]^2}{[1]^2}\right]e^{-tp}\\
&+\left[2\cos 3\theta\left(\frac{\mathbf{q}^{3p}[N]^2[N+1]^2[N+2]^2+\mathbf{q}^{-3p}[N]^2[N-1]^2[N-2]^2}{[1]^2[2]^2[3]^2}+\frac{[N+1]^2[N]^2[N-1]^2}{[1]^4[3]^2}\right)\right.\\
&+2\cos\theta\frac{\mathbf{q}^p[N+2]^2[N]^2[N-1]^2+\mathbf{q}^{-p}[N-2]^2[N]^2[N+1]^2}{[1]^4[2]^2}\\
&+2\cos(N+2)\theta\frac{(\mathbf{q}^{3p}+\mathbf{q}^{-p})[N]^2[N+1]^2}{[1]^2[2]^2}+2\cos(N-2)\theta\frac{(\mathbf{q}^{-3p}+\mathbf{q}^{p})[N]^2[N-1]^2}{[1]^2[2]^2}\\
&\left.+2\cos N\theta\frac{[N-1]^2[N+1]^2}{[1]^4}\right]e^{-\frac{3tp}{2}}+\mathcal{O}(e^{-2tp}).
\end{aligned}\tag{2.20}$$

As a consistency check of our result (2.20), we can take the limit

$$p\to\infty,\quad g=g_sp,\;\; A=gN:\text{fixed},\tag{2.21}$$

in which the partition function of qYM reduces to the partition function of undeformed Yang-Mills theory [6, 14, 15]. After taking this limit, the free energy becomes

$$\begin{aligned}
\log\frac{Z_N}{Z_{\rm gnd}} &= \left(1+\frac{A^2}{g^2}\right)\left[2e^{-\frac{A}{2}}-e^{-A}+\frac{8}{3}e^{-\frac{3A}{2}}\right]\\
&+\left[-\frac{2A^3}{g^3}\sinh g+\left(5+\frac{A^2}{g^2}\right)\frac{2A^2}{g^2}\sinh^2\frac{g}{2}\right]e^{-A}\\
&+\left[-\left(5+\frac{A^2}{g^2}\right)\frac{8A^3}{3g^3}\sinh^3 g+\left(13\cosh 2g+26\cosh g-3\right)\frac{4A^2}{9g^2}\sinh^2\frac{g}{2}\right.\\
&\left.+\left(11\cosh 2g+22\cosh g-15\right)\frac{8A^4}{9g^4}\sinh^2\frac{g}{2}+\left(\cosh g+2\right)\frac{16A^6}{9g^6}\sinh^4\frac{g}{2}\right]e^{-\frac{3A}{2}}+\mathcal{O}(e^{-2A}).
\end{aligned}\tag{2.22}$$

Here we have set $\theta=0$ for simplicity. This can be further expanded in the coupling g as

$$\log\frac{Z_N}{Z_{\rm gnd}}=\sum_{h=0}^{\infty}g^{2h-2}F_h(A),\tag{2.23}$$

where the first three terms read

$$
\begin{aligned}
F_0(A) &= 2A^2 e^{-\frac{A}{2}} + \left(-1-2A+\frac{A^2}{2}\right)A^2e^{-A} + \left(\frac{8}{3}+4A^2-\frac{8A^3}{3}+\frac{A^4}{3}\right)A^2e^{-\frac{3A}{2}} + \mathcal{O}(e^{-2A}),\\
F_1(A) &= 2e^{-\frac{A}{2}} + \left(-1+\frac{5A^2}{2}-\frac{A^3}{3}+\frac{A^4}{24}\right)e^{-A}\\
&\quad + \left(\frac{8}{3}+4A^2-\frac{40A^3}{3}+\frac{23A^4}{3}-\frac{4A^5}{3}+\frac{A^6}{9}\right)e^{-\frac{3A}{2}} + \mathcal{O}(e^{-2A}),\\
F_2(A) &= \left(\frac{5}{24}-\frac{A}{60}+\frac{A^2}{720}\right)A^2e^{-A} + \left(\frac{14}{3}-\frac{20A}{3}+\frac{221A^2}{90}-\frac{13A^3}{45}+\frac{13A^4}{720}\right)A^2e^{-\frac{3A}{2}} + \mathcal{O}(e^{-2A}).
\end{aligned} \tag{2.24}
$$

As expected, the above $F_0(A)$ agrees with the genus-zero free energy of undeformed Yang-Mills theory computed in [16].[3]

Chiral partition function. One can naturally distinguish the excitations as "chiral", "non-chiral", and "anti-chiral", as follows. In the chiral excitation, changes from the ground state are allowed only near the positive Fermi level $n = +n_F$. In other words, a chiral excitation is a configuration where the modes near the negative Fermi level are the same as the ground state

$$
\vec{n} = (-n_F, -n_F+1, \cdots, n_{N-1}, n_N). \tag{2.25}
$$

Figure 2(a) is an example of chiral excitation. The anti-chiral excitation is the momentum flip $n_i \to -n_i$ of chiral excitation, i.e. only the excitations near the negative Fermi level $n = -n_F$ are allowed. If the excitation involves both of the Fermi levels $n = \pm n_F$ it is called non-chiral (see figure 2(b) for an example of non-chiral excitation). This type of decomposition was first considered in the undeformed Yang-Mills theory in [24–26].

Now we can define the chiral partition function Z_N^+ by summing over the chiral excitations only. Z_N^+ is easily found from the full partition function (2.20) by dropping the non-chiral and anti-chiral terms

$$
\begin{aligned}
\frac{Z_N^+}{Z_{\rm gnd}} &= 1 + \frac{[N]^2}{[1]^2}e^{-\frac{Ng_sp}{2}+\mathrm{i}\theta} + \frac{\mathbf{q}^p[N]^2[N+1]^2+\mathbf{q}^{-p}[N]^2[N-1]^2}{[1]^2[2]^2}e^{-Ng_sp+2\mathrm{i}\theta}\\
&+ \left(\frac{\mathbf{q}^{3p}[N]^2[N+1]^2[N+2]^2+\mathbf{q}^{-3p}[N]^2[N-1]^2[N-2]^2}{[1]^2[2]^2[3]^2}+\frac{[N+1]^2[N]^2[N-1]^2}{[1]^4[3]^2}\right)e^{-\frac{3Ng_sp}{2}+3\mathrm{i}\theta}\\
&+ \mathcal{O}(e^{-2Ng_sp+4\mathrm{i}\theta}).
\end{aligned} \tag{2.26}
$$

It turns out that the chiral free energy $\log Z_N^+$ can be organized into a double expansion in terms of Q and $\widetilde{Q}$ in the 't Hooft limit, where Q and $\widetilde{Q}$ are defined by

$$
\begin{aligned}
Q &= e^{-T}, \qquad T = \frac{p-2}{2}Ng_s - \mathrm{i}\theta,\\
\widetilde{Q} &= e^{-t}, \qquad t = Ng_s.
\end{aligned} \tag{2.27}
$$

[3]Our $F_0(A)$ and that in [16] differ by a factor of A^2 which comes from the different definition of the genus expansion. In [16] the genus expansion is defined as the $1/N$ expansion, $\log Z = \sum_{h\geq 0} N^{2-2h}F_h(A)$, while we expand the free energy in terms of $g = A/N$.

The terms involving $\widetilde{Q}$ correspond to open string amplitudes due to the additional D-brane insertions on X_p [18]. The pure closed string amplitude is obtained by discarding the $\widetilde{Q}$ dependent terms from (2.26)

$$\frac{Z_N^{+,\text{closed}}}{Z_{\text{gnd}}} = 1+\frac{1}{[1]^2}Q+\frac{\mathbf{q}^{p-1}+\mathbf{q}^{-p+1}}{[1]^2[2]^2}Q^2+\left(\frac{\mathbf{q}^{3p-3}+\mathbf{q}^{-3p+3}}{[1]^2[2]^2[3]^2}+\frac{1}{[3]^2[1]^4}\right)Q^3+\mathcal{O}(Q^4), \quad (2.28)$$

and the closed string free energy $F^{\text{closed}} = \log(Z_N^{+,\text{closed}}/Z_{\text{gnd}})$ is given by

$$F^{\text{closed}} = \sum_{n=1}^{3}\frac{Q^n}{n[n]^2}+\frac{[p][p-2]}{[1]^2[2]^2}Q^2+\Big([p-1]^2+[1]^2+6\Big)\frac{[p][p-1]^2[p-2]}{[1]^2[2]^2[3]^2}Q^3+\mathcal{O}(Q^4). \quad (2.29)$$

One can show that the genus expansion of F^{closed} in (2.29) reproduces the result of topological string on X_p [27–29].

2.2 Weak coupling phase

Next consider the weak coupling phase ($t < t_c$). To study the weak coupling phase, we should perform the modular S-transformation of the Jacobi theta function in (2.9). Using the formula

$$\vartheta_{3-\epsilon}(v,\tau) = (-\mathrm{i}\tau)^{-\frac{1}{2}}e^{-\frac{\pi \mathrm{i} v^2}{\tau}}\vartheta_{3+\epsilon}\left(\frac{v}{\tau},-\frac{1}{\tau}\right), \quad (2.30)$$

the exact partition function in (2.9) becomes

$$Z_N = \left(\frac{2\pi}{g_s p}\right)^{\frac{N}{2}} e^{-\frac{N\theta^2}{2g_s p}} \det\left[e^{\frac{g_s}{2p}(i+j-N-1)^2}\vartheta_{3+\epsilon}\left(\frac{i+j-N-1}{p}-\frac{\mathrm{i}\theta}{g_s p},\frac{2\pi \mathrm{i}}{g_s p}\right)\right]. \quad (2.31)$$

Then, plugging the series expansion of Jacobi theta function

$$\vartheta_{3+\epsilon}\left(\frac{i+j-N-1}{p}-\frac{\mathrm{i}\theta}{g_s p},\frac{2\pi \mathrm{i}}{g_s p}\right) = \sum_{m\in\mathbb{Z}}(-1)^{\epsilon m}e^{-\frac{2\pi^2 m^2}{g_s p}+\frac{2\pi m\theta}{g_s p}}e^{\frac{2\pi \mathrm{i} m}{p}(i+j-N-1)} \quad (2.32)$$

into the determinant of (2.31), Z_N is written as a sum over integer vectors $\vec{m} = (m_1,\cdots,m_N)$

$$Z_N = \sum_{\vec{m}\in\mathbb{Z}^N}\Omega(\vec{m})\exp\left(-\frac{2\pi^2}{g_s p}\sum_{i=1}^{N}m_i^2+\frac{2\pi\theta}{g_s p}\sum_{i=1}^{N}m_i\right), \quad (2.33)$$

with some coefficient $\Omega(\vec{m})$. This S-dual expression has a natural interpretation as the instanton expansion. As in the case of undeformed Yang-Mills theory, the instanton in question is a classical solution of gauge field where the Dirac monopole configuration is embedded in the Cartan part of gauge field [30, 31]. In the weak coupling phase, the most dominant contribution comes from the zero-instanton sector since the instanton contribution $\vec{m} \neq 0$ in (2.33) is suppressed by the factor $\mathcal{O}(e^{-1/g_s})$ which is non-perturbative in g_s.

By setting $\vartheta_{3+\epsilon} = 1$ in (2.31) (or taking the $m = 0$ term in (2.32)), the perturbative part of Z_N in the weak coupling phase is found to be

$$\begin{aligned} Z_{\text{weak}} &= \left(\frac{2\pi}{g_s p}\right)^{\frac{N}{2}} e^{-\frac{N\theta^2}{2g_s p}} \det\left[e^{\frac{g_s}{2p}(i+j-N-1)^2}\right] \\ &= \left(\frac{2\pi}{g_s p}\right)^{\frac{N}{2}} e^{-\frac{N\theta^2}{2g_s p}+\frac{g_s(N^3-N)}{12p}} \prod_{k=1}^{N-1}\left(2\sinh\frac{g_s k}{2p}\right)^{N-k}. \end{aligned} \tag{2.34}$$

This is exactly the same as the partition function of pure Chern-Simons theory on S^3 up to a rescaling of the coupling $g_s \to g_s/p$. The genus expansion of (2.34) is easily found from the known result of pure Chern-Simons theory (see e.g. [32])

$$\begin{aligned} F_0(t) &= p^2\left[\frac{t^3}{6p^3} - \frac{\pi^2 t}{6p} - \mathrm{Li}_3(e^{-\frac{t}{p}}) + \zeta(3)\right], \\ F_1(t) &= -\frac{t}{8p} - \frac{1}{12}\log(1-e^{-\frac{t}{p}}) + \frac{t}{g_s}\log\frac{2\pi}{g_s} + \zeta'(-1) + \frac{1}{12}\log\frac{g_s}{p}, \\ F_{g\geq 2}(t) &= p^{2-2g}\frac{B_{2g}}{2g(2g-2)!}\left[\mathrm{Li}_{3-2g}(e^{-\frac{t}{p}}) + \frac{B_{2g-2}}{2g-2}\right]. \end{aligned} \tag{2.35}$$

2.3 Phase transition at $\theta = 0$

As shown in [6, 14, 15], the qYM on S^2 at $\theta = 0$ has a third order phase transition. The mechanism of the phase transition is essentially the same as the undeformed case found by Douglas and Kazakov [16] where the phase transition occurs when the eigenvalue density saturates the bound $\rho(h) \leq 1$.

As we have seen in the previous subsection, the zero-instanton sector of the weak coupling phase is described by the pure Chern-Simons theory. From the known eigenvalue density of Chern-Simons matrix model [32]

$$\rho(h) = \frac{p}{\pi}\arccos\left(e^{-\frac{t}{2p}}\cosh\frac{th}{2}\right) \tag{2.36}$$

the critical point $t = t_c$ is determined by the condition $\rho(0) = 1$, i.e.

$$\frac{p}{\pi}\arccos\left(e^{-\frac{t_c}{2p}}\right) = 1, \tag{2.37}$$

which leads to the result (2.12) found in [6, 14, 15].

Now we can numerically study the behavior of free energy at $\theta = 0$ using the exact partition function (2.9) at finite N. The determinant in (2.9) can be evaluated numerically with high precision and we can plot the free energy as a function of $t = g_s N$ by varying the coupling g_s with fixed N. In figure 3 we show the plot of free energy for $p = 3, N = 80$ at $\theta = 0$. As we can see from this figure, the behavior of the free energy changes at $t = t_c$ from Z_{weak} (2.34) in the weak coupling phase to Z_{gnd} (2.17) in the strong coupling phase, as expected from the result in [6, 14, 15].

JHEP04(2018)059

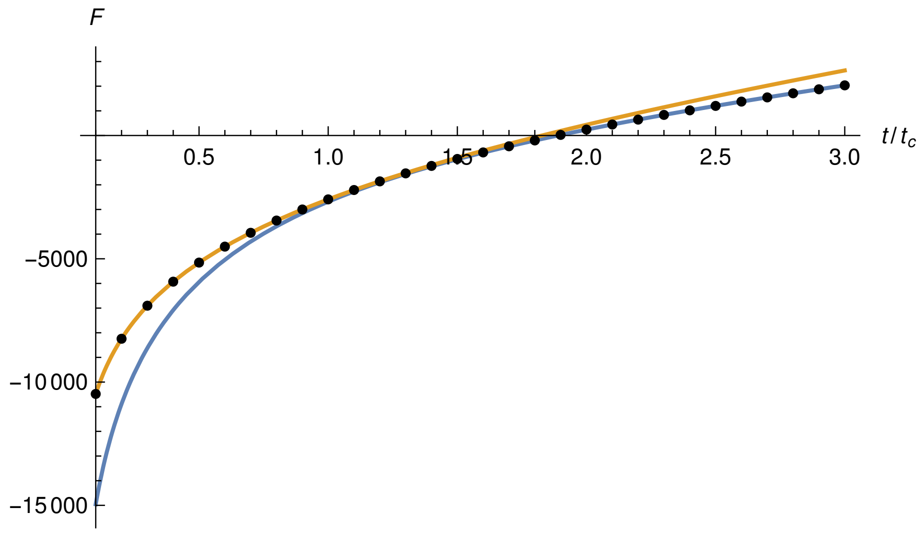


Figure 3. Plot of free energy $F = \log Z_N$ for $p = 3, N = 80$ at $\theta = 0$. The dots are the numerical values of the exact partition function (2.9), while the orange curve and the blue curve represent $\log Z_{\text{weak}}$ in (2.34) and $\log Z_{\text{gnd}}$ in (2.17), respectively.

3 Instanton in the weak coupling phase

It is argued in [6, 14, 15] that the phase transition of qYM is induced by instantons. As we have seen in the previous section 2.2, the instanton expansion in the weak coupling phase naturally arises after performing the modular S-transformation of Jacobi theta function (2.33).

It turns out that when $0 < \theta < \pi$ the dominant contribution comes from the $m = 0$ and $m = 1$ terms in the expansion of Jacobi theta function (2.32). If we keep only those terms in the expression of Z_N in (2.31), the partition function is approximated as

$$Z_N \approx \mathcal{N} \det(A + \xi B), \tag{3.1}$$

where ξ is the weight factor of 1-instanton

$$\xi = e^{\frac{2\pi\theta - 2\pi^2}{g_s p}}, \tag{3.2}$$

$\mathcal{N}$ is the overall factor

$$\mathcal{N} = \left(\frac{2\pi}{g_s p}\right)^{\frac{N}{2}} e^{-\frac{N\theta^2}{2g_s p}}, \tag{3.3}$$

and A and B in (3.1) are the following $N \times N$ matrices:

$$\begin{aligned} A_{i,j} &= q^{\frac{1}{2}(i+j-N-1)^2}, \\ B_{i,j} &= (-1)^{N-1} x^{i+j-N-1} A_{i,j}, \qquad (i, j = 1, \cdots, N). \end{aligned} \tag{3.4}$$

Here we have introduced the notation q and x by

$$q = e^{-\frac{g_s}{p}}, \quad x = e^{\frac{2\pi \mathrm{i}}{p}}. \tag{3.5}$$

The validity of this approximation (3.1) will be discussed in detail in the next section 4. Note that the sign $(-1)^{N-1} = (-1)^{\epsilon}$ of $B_{i,j}$ in (3.4) comes from the sign $(-1)^{\epsilon m}$ in the expansion of Jacobi theta function (2.32) with $m = 1$. In this notation, Z_{weak} in (2.34) is written as

$$Z_{\text{weak}} = \mathcal{N} \det A. \tag{3.6}$$

Now we can define the instanton part of partition function by dividing Z_N by Z_{weak}. In the approximation (3.1) the instanton partition function becomes

$$Z_{\text{inst}} = \frac{\det(A + \xi B)}{\det A} = \det(1 + \xi M), \tag{3.7}$$

where the matrix M is given by

$$M = A^{-1} B. \tag{3.8}$$

From the explicit form of A and B in (3.4), we find that the matrix element of M has a simple expression

$$M_{i,j} = (-1)^{N-1+j-i} \frac{(x^{-1}q;q)_{i-1}}{(q;q)_{i-1}} \frac{(xq;q)_{N-j}}{(q;q)_{N-j}} \frac{x-1}{x-q^{j-i}} x^{2j-N-1} q^{\frac{1}{2}(j-i)(N+2-i-j)}, \tag{3.9}$$

where $(a;q)_k$ denotes the q-Pochhammer symbol

$$(a;q)_k = \prod_{n=0}^{k-1} (1 - aq^n). \tag{3.10}$$

We have checked this relation (3.9) for $N \le 10$, and we believe that this is true for all N. In what follows we will assume that (3.9) holds for all N. It would be interesting to find a general proof of (3.9).

One can expand Z_{inst} in (3.7) as a power series in ξ

$$Z_{\text{inst}} = \sum_{k=0}^{\infty} Z_k, \tag{3.11}$$

where $Z_0 = 1$ and $Z_k \propto \xi^k$. For instance, the 1-instanton term is given by

$$Z_1 = \xi \operatorname{Tr} M. \tag{3.12}$$

Higher instanton corrections $Z_{k\ge 2}$ will be studied in section 4. In the 't Hooft limit, we expect that the 1-instanton correction Z_1 at $\theta = 0$ is characterized by the instanton action $S_{\text{inst}}(t)$ computed in [6, 14, 15]

$$\xi_0 \operatorname{Tr} M \sim e^{-\frac{1}{g_s} S_{\text{inst}}(t)}, \tag{3.13}$$

with ξ_0 being

$$\xi_0 = \xi_{\theta=0} = e^{-\frac{2\pi^2}{g_s p}}. \tag{3.14}$$

JHEP04(2018)059

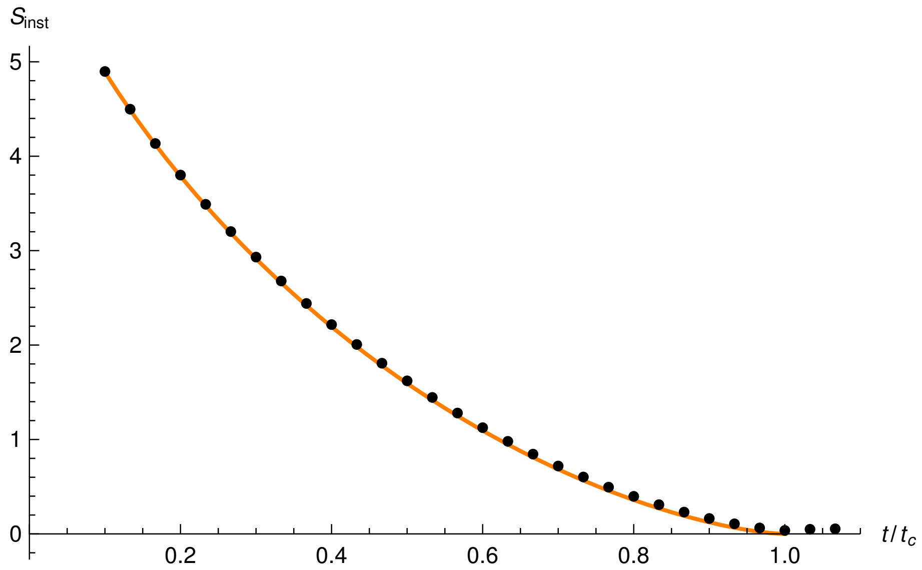


Figure 4. Plot of instanton action at $\theta = 0$. The dots are the numerical value of $-g_s \log(\xi_0 \operatorname{Tr} M)$ for $p = 3, N = 400$, while the orange curve represents $S_{\text{inst}}(t)$ in (3.15).

In [6], it was found that the instanton action is given by the integral of eigenvalue density $\rho(h)$ in (2.36) along the *imaginary axis*

$$S_{\text{inst}}(t) = 2\pi t \int_0^{h_0} dh \Big[1 - \rho(\mathrm{i}h)\Big]. \tag{3.15}$$

Here the upper bound of integral, h_0, is determined by the condition $\rho(\mathrm{i}h_0) = 1$:

$$\rho(\mathrm{i}h) = \frac{p}{\pi} \arccos\left(e^{-\frac{t}{2p}} \cos\frac{th}{2}\right), \quad h_0 = \frac{2}{t} \arccos\left(e^{\frac{t}{2p}} \cos\frac{\pi}{p}\right). \tag{3.16}$$

This suggests that the large N limit of instanton in the weak coupling phase can be thought of as a complex instanton. A similar phenomenon was observed in the GWW model as well [11, 33].

In figure 4 we show the plot of $\xi_0 \operatorname{Tr} M$ for $p = 3, N = 400$ using the exact form of M in (3.9). One can clearly see that the exact result of M nicely reproduces the analytic form of instanton action in (3.15). The instanton action vanishes at $t = t_c$ as shown in [6, 14, 15], which is also reproduced numerically by our exact result of M. This leads to a physical picture of the phase transition that it is triggered by the condensation of instantons, as in the case of undeformed theory [17]. We also observed numerically that the 1-instanton correction $\xi_0 \operatorname{Tr} M$ is always positive for both even N and odd N in the weak coupling phase; we emphasize that the sign $(-1)^{N-1}$ of B in (3.4) is crucial for this positivity of 1-instanton correction.

3.1 Z_{inst} in the $p \to \infty$ limit

We expect that Z_{inst} reduces to the known instanton correction of undeformed theory in the limit $p \to \infty$ with $g = g_s p$ fixed (2.21). Indeed, we find that the 1-instanton term (3.12) reduces to

$$\lim_{p\to\infty} \operatorname{Tr} M = (-1)^{N-1} L^1_{N-1}\left(\frac{4\pi^2}{g}\right), \tag{3.17}$$

where $L^\alpha_n(x) = \frac{1}{n!}x^{-\alpha}e^x\partial^n_x(x^{n+\alpha}e^{-x})$ denotes the Laguerre polynomial. As expected, this agrees with the result of 1-instanton correction of undeformed theory [17]. This implies that the 1-instanton term $\operatorname{Tr} M$ in qYM can be thought of as a certain q-deformation of the Laguerre polynomial. As shown in [17], the large N limit of $L^1_{N-1}(4\pi^2/g)$ has a sign $(-1)^{N-1}$ which is precisely canceled by the overall sign in (3.17) coming from the modular S-transformation of Jacobi theta function. The resulting 1-instanton factor (3.17) in the undeformed theory is always positive for both even N and odd N, which is consistent with the above observation of the positivity of $\operatorname{Tr} M$ in the q-deformed theory.

More generally, we find

$$\lim_{p\to\infty} \det(1+\xi M) = \det(1+\xi\widetilde{M}), \tag{3.18}$$

where the $N \times N$ matrix $\widetilde{M}$ is given by

$$\widetilde{M}_{i,j} = (-1)^{N-1} L^{j-i}_{i-1}\left(\frac{4\pi^2}{g}\right), \quad (i,j=1,\cdots,N). \tag{3.19}$$

The right hand side of (3.18) is exactly the generating function of the expectation value of 't Hooft loops in the anti-symmetric representations in 4d $\mathcal{N}=4$ U(N) super Yang-Mills theory,[4] up to a change of sign of the coupling $g = -g^2_{4d}$ [34]. This is expected since the computation of 't Hooft loops in 4d $\mathcal{N}=4$ super Yang-Mills theory localizes to 2d Yang-Mills theory on S^2 in the instanton sector [34].

We have checked that (3.17) and (3.18) hold for small N, but we do not have a general proof. It would be nice to find a proof of (3.17) and (3.18) for general N.

4 Phase diagram of q-deformed Yang-Mills theory at $\theta \neq 0$

The phase diagram of qYM at non-zero θ was conjectured in [6]. In this section we will examine this conjecture using our exact result of partition function at finite N.

We first notice that using the symmetry of the exact partition function

$$\theta \to -\theta, \qquad \theta \to \theta + 2\pi, \tag{4.1}$$

we can restrict θ to the region $0 \le \theta \le \pi$ without loss of generality. As discussed in [36], for general value of θ we should minimize the energy of Z_{weak} in (2.34) under all 2π-shifts of θ

$$\min_{\ell\in\mathbb{Z}} \frac{N(\theta+2\pi\ell)^2}{2g_s p}. \tag{4.2}$$

[4] This generating function of 't Hooft loops is obtained from the exact result of Wilson loops in [35] by the S-duality of 4d $\mathcal{N}=4$ super Yang-Mills theory.

JHEP04(2018)059

The minimum is given by $\ell = 0$ if θ is in the range $0 \le \theta \le \pi$, and hence we can safely use the S-dual expression of Z_N in (2.31) and (3.1) in this region of θ. Thus the 1-instanton term for $0 \le \theta \le \pi$ is given by

$$Z_1 = \xi \operatorname{Tr} M = e^{\frac{2\pi\theta}{g_s p}} \xi_0 \operatorname{Tr} M \sim e^{\frac{2\pi\theta}{g_s p} - \frac{1}{g_s} S_{\text{inst}}(t)}, \tag{4.3}$$

where we used (3.13). Then the critical value $t = t_*(\theta)$ where Z_1 becomes of order one is determined by the condition that the exponent in (4.3) vanishes

$$S_{\text{inst}}\big(t_*(\theta)\big) = \frac{2\pi\theta}{p}. \tag{4.4}$$

It is conjectured in [6] that the critical line $t = t_*(\theta)$ on the t-θ plane is just the first one of such critical lines; there are many critical lines on the t-θ plane which accumulate at $(t, \theta) = (0, \pi)$ (see figure 3 in [6]).

The conjectured phase diagram of [6] is based on the following two assumptions:

(i) Only the instantons with charges $\vec{m} = (1, 1, \cdots, 1, 0 \cdots, 0)$ in the expansion (2.33) are relevant for the phase transition at $\theta \neq 0$.

(ii) There is a series of phase transitions at $\theta \neq 0$ where the instantons of the type in (i) exchange dominance.

The assumption (i) amounts to using the approximation in (3.1). We can test this assumption by computing the following ratio numerically

$$r = \frac{Z_{\text{weak}} Z_{\text{inst}}}{Z_N}, \tag{4.5}$$

where Z_{weak} and Z_{inst} are given by (2.34) and (3.7), respectively. In figure 5, we show the plot of this ratio for $p = 3, N = 80$ at $\theta = \frac{\pi}{3}$. From this figure, one can see that this ratio is very close to 1

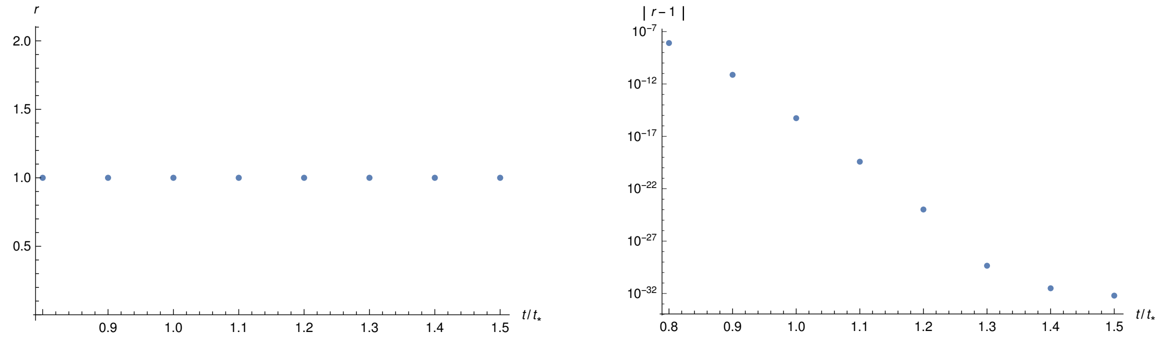
$$r \approx 1, \tag{4.6}$$

which confirms the assumption (i).

Next consider the assumption (ii). In figure 6, we show the plot of the instanton part of free energy $F_{\text{inst}} = \log Z_{\text{inst}}$ and its derivatives for $p = 3, N = 80$ at $\theta = \frac{\pi}{3}$.[5] One can see that the third derivative $\partial_t^3 F_{\text{inst}}$ in figure 6d has several jumps at different values of t, and the first jump (or discontinuity of $\partial_t^3 F_{\text{inst}}$) occurs at $t \approx t_*(\theta)$. This is consistent with the assumption (ii). In figure 6 we have used the approximate instanton partition function Z_{inst} in (3.7), but the plot of the exact partition function Z_N/Z_{weak} does not change much from figure 6 due to the property $r \approx 1$ (4.6). In figure 6d, the third derivative $\partial_t^3 F_{\text{inst}}$ has sharp discontinuities only at the first few zeros of $\partial_t^3 F_{\text{inst}}$, but we expect that this is a finite N effect and in the strict large N limit they become sharp phase transitions.

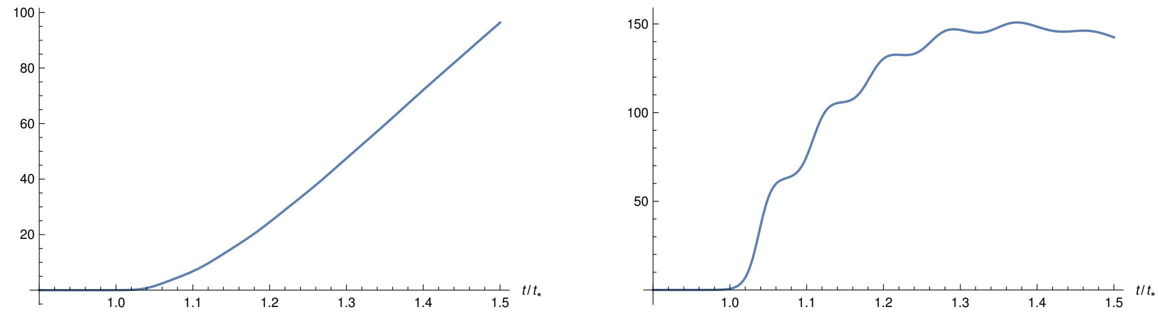
We can collect more evidence for the assumption (ii) by evaluating the k-instanton contribution Z_k in (3.11) separately. In figure 7, we show the plot of the k-instanton

[5] To draw this plot, we first compute $F_{\text{inst}}(t)$ numerically at discrete values of t, and then find the interpolating function from the discrete data. In figure 6b–6d, we plot the derivatives of this interpolating function.

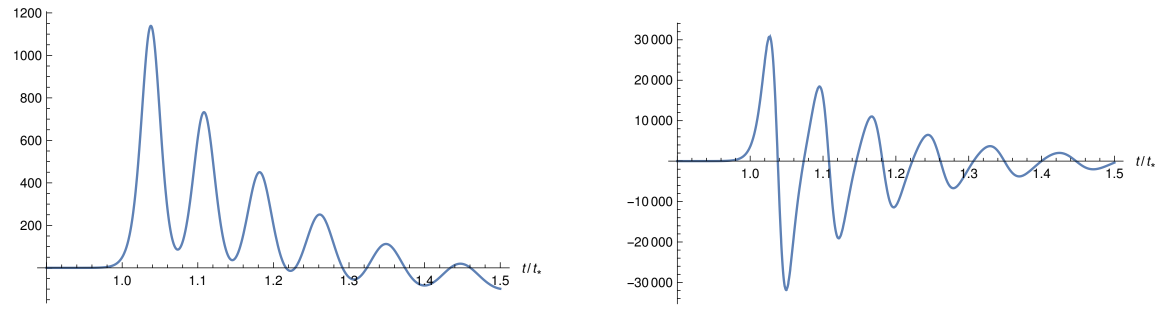


(a) Plot of r. (b) Plot of $|r-1|$.

Figure 5. Plot of the ratio r (4.5) for $p=3, N=80$ at $\theta=\frac{\pi}{3}$. The horizontal axis is $t/t_*(\pi/3)$ where $t_*(\theta)$ is defined by (4.4).



(a) F_{inst}. (b) $\partial_t F_{\text{inst}}$.

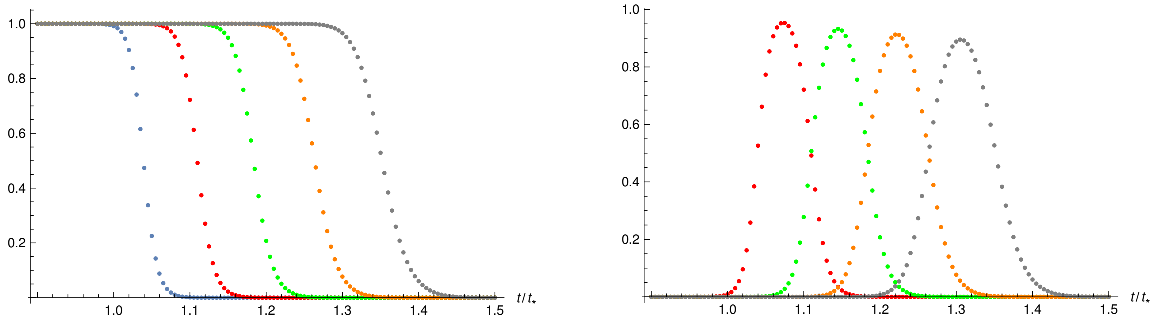


(c) $\partial_t^2 F_{\text{inst}}$. (d) $\partial_t^3 F_{\text{inst}}$.

Figure 6. Plot of $F_{\text{inst}} = \log Z_{\text{inst}}$ and its derivatives for $p=3, N=80$ at $\theta=\frac{\pi}{3}$.

contribution ($k=1,\cdots,4$) for $p=3, N=80$ at $\theta=\frac{\pi}{3}$. One can see that the dominant instanton changes from $k=1$ to $k=4$ as t increases, and the exchange of dominance occurs at different values of t for different instanton number k. In other words, there is a series of phase transitions at $t=t_k$ $(k=1,2,\cdots)$ where Z_{k-1} and Z_k exchange dominance at $t=t_k$. Our numerical result in figure 7 gives strong evidence for the assumption (ii). Also, we observed numerically that the difference of t_k and t_{k+1} decreases as N becomes large, and the difference scales approximately as $1/N$

$$t_{k+1} - t_k \sim 1/N. \tag{4.7}$$



(a) Plot of $\frac{\sum_{n=0}^{k} Z_n}{Z_{\text{inst}}}$, $(k = 0, \cdots, 4)$.

(b) Plot of $\frac{Z_k}{Z_{\text{inst}}}$, $(k = 1, \cdots, 4)$.

Figure 7. Plot of k-instanton contribution for $p = 3, N = 80$ at $\theta = \frac{\pi}{3}$. In a we show the sum of instanton contributions up to k-instantons normalized by Z_{inst}, and in b we show each k-instanton contribution individually. In both figures a and b, we used the same colors: $k = 1$ (red), $k = 2$ (green), $k = 3$ (orange), $k = 4$ (gray).

This suggests that the $1/N$ correction of instanton factor, in particular the *prefactor* f_k of k-instanton, is important for the understanding of the scaling behavior (4.7)

$$Z_k = f_k(t, g_s) e^{\frac{2\pi\theta k}{g_s p} - \frac{k}{g_s} S_{\text{inst}}(t)}. \tag{4.8}$$

However, we were unable to find the analytic form of the prefactor f_k and hence we could not determine the analytic form of the critical value $t = t_k$. In the next section, we will consider the phase diagram of undeformed Yang-Mills theory, where the instanton prefactor is more tractable analytically.

Finally, we can draw the phase diagram of qYM by using our exact result of Z_{inst} in (3.7). To do this, we first observe from figure 6 that the local maximum of the second derivative $\partial_t^2 F_{\text{inst}}$ corresponds to the (approximate) discontinuous point of the third derivative $\partial_t^3 F_{\text{inst}}$. Based on this observation, in figure 8 we plot the local maxima of $\partial_t^2 F_{\text{inst}}$ for $p = 3, N = 80$ for several values of θ. Our result in figure 8 agrees with the conjectured phase diagram in [6], at least qualitatively. In particular we can see from figure 8 that the transition curves seem to accumulate at the point $(t, \theta) = (0, \pi)$.

5 Phase diagram of undeformed Yang-Mills theory at $\theta \neq 0$

In this section, we consider the phase transition curves in the undeformed theory, where the instanton prefactor can be studied analytically. In fact, the analytic form of the prefactor of 1-instanton term has been already obtained in [17].

Let us consider the instanton contributions in the undeformed Yang-Mills theory (3.18)

$$Z_{\text{inst}} = \det(1 + \xi \widetilde{M}) \tag{5.1}$$

where $\widetilde{M}$ is given by (3.19). The coupling in the undeformed theory is defined in (2.21). It is convenient to rescale the coupling g and A in (2.21) as

$$\hat{g} = \frac{g}{\pi^2}, \quad a = \hat{g}N = \frac{A}{\pi^2}. \tag{5.2}$$

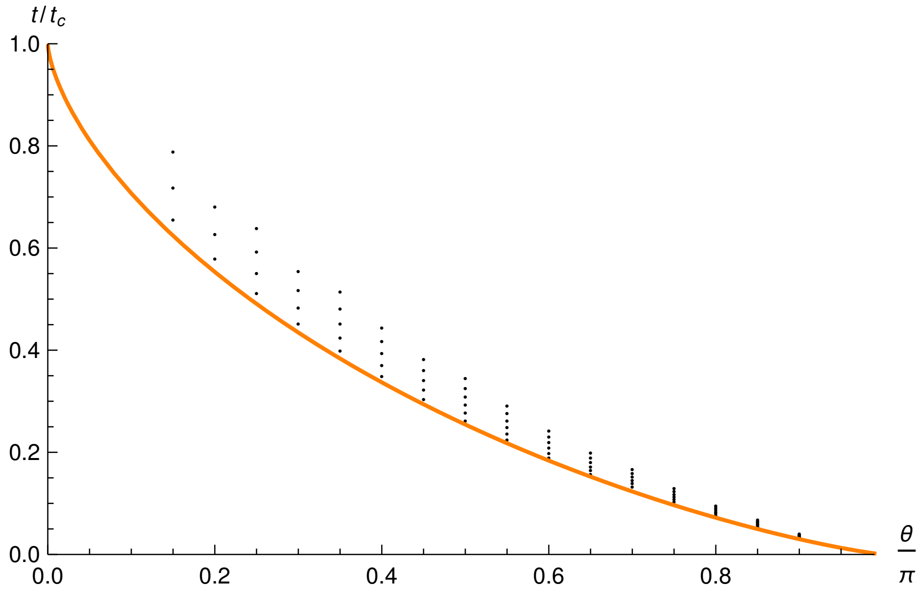


Figure 8. Phase diagram at non-zero θ ($p = 3, N = 80$). We plotted the local maxima of $\partial_t^2 F_{\text{inst}}$ at fixed θ in the range $t_*(\theta) \leq t \leq 1.4t_*(\theta)$ and varied θ with step $\Delta\theta/\pi = 0.05$. The orange curve is the line $t = t_*(\theta)$ given by (4.4).

In this normalization, when $\theta = 0$ the phase transition occurs at $a = 1$, corresponding to the critical value $A = \pi^2$ found in [16]. In terms of these rescaled couplings, $\widetilde{M}$ and ξ in (5.1) become

$$\widetilde{M}_{i,j} = (-1)^{N-1} L_{i-1}^{j-i}(4/\hat{g}), \qquad \xi = e^{\frac{2(\theta-\pi)}{\pi\hat{g}}}. \tag{5.3}$$

By expanding Z_{inst} in (5.1) as a power series in ξ we can define the k-instanton term Z_k as in the case of qYM in (3.11). As we have seen in (3.17), the 1-instanton term can be written in a closed form

$$Z_1 = \xi \operatorname{Tr} \widetilde{M} = \xi(-1)^{N-1} L_{N-1}^1(4/\hat{g}). \tag{5.4}$$

The large N limit of Z_1 has been studied in [17] and the result reads

$$Z_1 = f_1(a, \hat{g}) e^{-\frac{1}{\hat{g}} S_{\text{inst}}(a) + \frac{2\theta}{\pi\hat{g}}}, \tag{5.5}$$

where the instanton action is given by[6]

$$S_{\text{inst}}(a) = 2\left[\sqrt{1-a} - a\cosh^{-1}(1/\sqrt{a})\right]. \tag{5.7}$$

[6]This is obtained by integrating the eigenvalue density of Gaussian matrix model along the imaginary axis, as in the case of q-deformed theory (3.15)

$$S_{\text{inst}}(a) = 2\pi a \int_0^{h_0} dh \big[1 - \rho_{\text{G}}(\mathrm{i}h)\big], \quad \rho_{\text{G}}(\mathrm{i}h_0) = 1, \quad \rho_{\text{G}}(h) = \frac{\pi a}{2}\sqrt{\frac{4}{\pi^2 a} - h^2}. \tag{5.6}$$

The prefactor $f_1(a,\hat{g})$ was also computed in [17][7]

$$f_1(a,\hat{g})=\frac{1}{4}\left[\frac{\hat{g}^2a^2}{4\pi^2(1-a)}\right]^{\frac{1}{4}}\Big(1+\mathcal{O}(\hat{g})\Big). \tag{5.8}$$

As we mentioned in section 3, the sign $(-1)^{N-1}$ in (5.4) is precisely canceled by the same sign coming from the large N limit of Laguerre polynomial, and the final result of the prefactor $f_1(a,\hat{g})$ does not have this sign. We expect that the k-instanton contribution Z_k has the form

$$Z_k=f_k(a,\hat{g})e^{-\frac{k}{\hat{g}}S_{\text{inst}}(a)+\frac{2k\theta}{\pi\hat{g}}}. \tag{5.9}$$

Namely, $\log Z_k\approx k\log Z_1$ at the leading order in $\hat{g}$ expansion. We have checked this behavior numerically for $k=2,3,4$. We would like to find the prefactor $f_k(a,\hat{g})$ for $k\geq 2$ but we were unable to determine them analytically. However, one can study the prefactor $f_k(a,\hat{g})$ numerically using the exact result at finite N. For instance, from the numerical analysis of the large N behavior of the 2-instanton

$$Z_2=\frac{\xi^2}{2}\left[\Big(\operatorname{Tr}\widetilde{M}\Big)^2-\operatorname{Tr}\Big(\widetilde{M}\Big)^2\right], \tag{5.10}$$

and assuming $f_2(a,\hat{g})=\alpha f_1(a,\hat{g})^\beta$ with some constants α,β at the leading order in $\hat{g}$ expansion, we can determine the parameters α,β numerically. In this way we find the prefactor of 2-instanton

$$f_2(a,\hat{g})=\frac{\pi}{128}\frac{\hat{g}^2a^2}{4\pi^2(1-a)}\Big(1+\mathcal{O}(\hat{g})\Big). \tag{5.11}$$

It would be interesting to derive this result analytically.

These prefactors are important to find the critical lines at non-zero θ. The critical value $a=a_*(\theta)$ at the leading order in $\hat{g}$ expansion is determined by the condition that the exponent of Z_k in (5.9) vanishes

$$S_{\text{inst}}\big(a_*(\theta)\big)=\frac{2\theta}{\pi}. \tag{5.12}$$

This vanishing condition of the exponential factor is common for all k, and this condition alone is not enough to distinguish the dominant Z_k. It turns out that it is important to include the effect of prefactor to explain the splitting of critical values observed in (4.7). Let us consider the $\hat{g}$ correction for the first two critical values $a_1(\theta),a_2(\theta)$ determined by the condition

$$\frac{Z_1}{Z_0}(a_1(\theta))=1,\quad \frac{Z_2}{Z_1}(a_2(\theta))=1. \tag{5.13}$$

Including the contribution of prefactors, we find the $\mathcal{O}(\hat{g})$ deviation of $a_1(\theta)$ and $a_2(\theta)$ from the leading term $a_*(\theta)$ in (5.12)

$$\begin{aligned}a_1(\theta)&=a_*(\theta)+\hat{g}\frac{\log f_1(a_*(\theta),\hat{g})}{S'_{\text{inst}}(a_*(\theta))}+\mathcal{O}(\hat{g}^2),\\ a_2(\theta)&=a_*(\theta)+\hat{g}\frac{\log f_2(a_*(\theta),\hat{g})-\log f_1(a_*(\theta),\hat{g})}{S'_{\text{inst}}(a_*(\theta))}+\mathcal{O}(\hat{g}^2),\end{aligned} \tag{5.14}$$

[7] It would be possible to compute the $\mathcal{O}(\hat{g})$ correction of $f_1(a,\hat{g})$ using the known asymptotic behavior of Laguerre polynomials [37].

JHEP04(2018)059

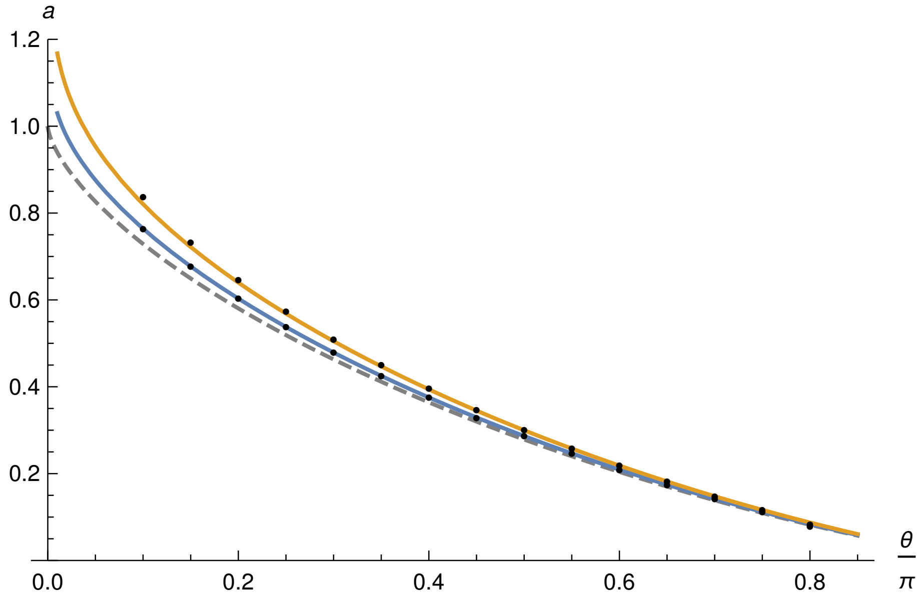


Figure 9. The phase transition lines of undeformed Yang-Mills theory at non-zero θ ($N = 80$). The black dots are the first two local maxima of $\partial_t^2 F_{\text{inst}}$ at fixed θ and we varied θ with step $\Delta\theta/\pi = 0.05$. The blue and the orange curves are the lines $a = a_1(\theta)$ and $a = a_2(\theta)$, respectively. The gray dashed curve represents the curve $a = a_*(\theta)$ without taking into account the effect of instanton prefactors.

where $S'_{\text{inst}}(a) = \partial_a S_{\text{inst}}(a)$ is given by

$$S'_{\text{inst}}(a) = -2\cosh^{-1}(1/\sqrt{a}). \tag{5.15}$$

This nicely explains the splitting $a_2 - a_1 \sim \mathcal{O}(1/N)$ observed in (4.7).

Now we can draw the phase diagram of undeformed theory in a similar manner as the q-deformed case in figure 8. In figure 9, we plot the first two maxima of $\partial_t^2 F_{\text{inst}}$ computed numerically from Z_{inst} in (5.1). One can see that numerical data points fit well on the curves $a = a_1(\theta)$ and $a = a_2(\theta)$ obtained in (5.14). It would also be interesting to study the transition curves $a = a_k(\theta)$ for higher instanton corrections $Z_{k\geq 3}$, which we will leave for a future problem.

6 Conclusion and open problems

In this paper we have examined the phase diagram of q-deformed Yang-Mills theory on S^2 conjectured in [6] and found numerical evidence for this conjecture using the exact partition function at finite N (2.9). We found that the $1/N$ correction to the instanton contribution, in particular the prefactor of instanton, is important for the understanding of the splitting of phase transition curves at non-zero θ. Our analysis heavily relied on numerics and it is desirable to find a more analytic method to study the phase diagram.

There are various open problems. We have seen that the instanton correction in qYM has an interesting connection to a q-deformation of Laguerre polynomials, in the sense

JHEP04(2018)059

that the 1-instanton term $Z_1 = \xi \operatorname{Tr} M$ reduces to the ordinary Laguerre polynomial in the limit (3.17). More generally, we observed that the characteristic polynomial of M in (3.9) reduces to that of $\widetilde{M}$ in (3.19) in this limit (3.18). However, at present it is not clear to us how the instanton correction in qYM is related to the standard definition of q-Laguerre polynomials, also known as the generalized Stieltjes-Wigert polynomials (see e.g. [38]). Understanding such a relation could be a key towards an analytic proof of the conjectured form of M in (3.9). Also, it is important to find the analytic form of the instanton prefactor $f_k(t, g_s)$ in the q-deformed case which plays an important role for the phase diagram at non-zero θ. We observed numerically that only the anti-symmetric representations of instantons are relevant and other representations are suppressed in the large N limit (4.6), as conjectured in [6]. It would be interesting to understand the physical origin of this phenomenon. It would also be important to understand the implication of this phase structure for the black hole physics. According to [18, 19] the coefficient $\Omega(\vec{m})$ in (2.33) is related to the black hole entropy, and it would be very interesting to study its large N behavior. It is argued in [19, 39, 40] that the non-perturbative $\mathcal{O}(e^{-N})$ effect is responsible for the failure of chiral factorization of partition function and it has an interesting consequence in the dual spacetime picture. It would be very interesting to study such $\mathcal{O}(e^{-N})$ effects in the q-deformed Yang-Mills theory from the viewpoint of resurgence, in a similar manner as the GWW model studied in [11, 13]. Some progress in this direction for the 2d Yang-Mills theory on a torus will be reported elsewhere [41]. Also, we expect that the chiral partition function of qYM receives "membrane instanton corrections" given by the Nekrasov-Shatashvili limit of the refined topological string on X_p from the general argument in [10]. Note that it was shown in [42] that the refined topological string on X_p is related to a two parameter (q, t)-deformation of 2d Yang-Mills theory and the large N phase structure of (q, t)-deformed 2d Yang-Mills on S^2 was studied in [43]. It would be very interesting to investigate this direction further.

Acknowledgments

I would like to thank Daniel Jafferis for correspondence. This work was supported in part by JSPS KAKENHI Grant No. 16K05316.

References

[1] C.G. Callan Jr., R.F. Dashen and D.J. Gross, *The structure of the gauge theory vacuum*, *Phys. Lett.* **B 63** (1976) 334 [INSPIRE].

[2] S.R. Coleman, *More about the massive Schwinger model*, *Annals Phys.* **101** (1976) 239 [INSPIRE].

[3] A. D'Adda, M. Lüscher and P. Di Vecchia, *A 1/n expandable series of nonlinear σ-models with instantons*, *Nucl. Phys.* **B 146** (1978) 63 [INSPIRE].

[4] E. Witten, *Phases of $N = 2$ theories in two-dimensions*, *Nucl. Phys.* **B 403** (1993) 159 [hep-th/9301042] [INSPIRE].

[5] D. Gaiotto, A. Kapustin, Z. Komargodski and N. Seiberg, *Theta, time reversal and temperature*, *JHEP* **05** (2017) 091 [arXiv:1703.00501] [INSPIRE].

[6] D. Jafferis and J. Marsano, *A DK phase transition in q-deformed Yang-Mills on S^2 and topological strings*, hep-th/0509004 [INSPIRE].

[7] Y. Hatsuda, S. Moriyama and K. Okuyama, *Exact results on the ABJM Fermi gas*, *JHEP* **10** (2012) 020 [arXiv:1207.4283] [INSPIRE].

[8] Y. Hatsuda, S. Moriyama and K. Okuyama, *Instanton effects in ABJM theory from Fermi gas approach*, *JHEP* **01** (2013) 158 [arXiv:1211.1251] [INSPIRE].

[9] Y. Hatsuda, S. Moriyama and K. Okuyama, *Instanton bound states in ABJM theory*, *JHEP* **05** (2013) 054 [arXiv:1301.5184] [INSPIRE].

[10] Y. Hatsuda, M. Mariño, S. Moriyama and K. Okuyama, *Non-perturbative effects and the refined topological string*, *JHEP* **09** (2014) 168 [arXiv:1306.1734] [INSPIRE].

[11] M. Mariño, *Nonperturbative effects and nonperturbative definitions in matrix models and topological strings*, *JHEP* **12** (2008) 114 [arXiv:0805.3033] [INSPIRE].

[12] K. Okuyama, *Wilson loops in unitary matrix models at finite N*, *JHEP* **07** (2017) 030 [arXiv:1705.06542] [INSPIRE].

[13] A. Ahmed and G.V. Dunne, *Transmutation of a trans-series: the Gross-Witten-Wadia phase transition*, *JHEP* **11** (2017) 054 [arXiv:1710.01812] [INSPIRE].

[14] X. Arsiwalla, R. Boels, M. Mariño and A. Sinkovics, *Phase transitions in q-deformed 2D Yang-Mills theory and topological strings*, *Phys. Rev.* **D 73** (2006) 026005 [hep-th/0509002] [INSPIRE].

[15] N. Caporaso, M. Cirafici, L. Griguolo, S. Pasquetti, D. Seminara and R.J. Szabo, *Topological strings and large N phase transitions. I. Nonchiral expansion of q-deformed Yang-Mills theory*, *JHEP* **01** (2006) 035 [hep-th/0509041] [INSPIRE].

[16] M.R. Douglas and V.A. Kazakov, *Large N phase transition in continuum QCD in two-dimensions*, *Phys. Lett.* **B 319** (1993) 219 [hep-th/9305047] [INSPIRE].

[17] D.J. Gross and A. Matytsin, *Instanton induced large N phase transitions in two-dimensional and four-dimensional QCD*, *Nucl. Phys.* **B 429** (1994) 50 [hep-th/9404004] [INSPIRE].

[18] M. Aganagic, H. Ooguri, N. Saulina and C. Vafa, *Black holes, q-deformed 2d Yang-Mills and non-perturbative topological strings*, *Nucl. Phys.* **B 715** (2005) 304 [hep-th/0411280] [INSPIRE].

[19] C. Vafa, *Two dimensional Yang-Mills, black holes and topological strings*, hep-th/0406058 [INSPIRE].

[20] H. Ooguri, A. Strominger and C. Vafa, *Black hole attractors and the topological string*, *Phys. Rev.* **D 70** (2004) 106007 [hep-th/0405146] [INSPIRE].

[21] A.A. Migdal, *Recursion equations in gauge theories*, *Sov. Phys. JETP* **42** (1975) 413 [*Zh. Eksp. Teor. Fiz.* **69** (1975) 810] [INSPIRE].

[22] D.J. Gross and E. Witten, *Possible third order phase transition in the large N lattice gauge theory*, *Phys. Rev.* **D 21** (1980) 446 [INSPIRE].

[23] S.R. Wadia, *A study of $\mathrm{U}(N)$ lattice gauge theory in 2-dimensions*, arXiv:1212.2906 [INSPIRE].

JHEP04(2018)059

[24] D.J. Gross, *Two-dimensional QCD as a string theory*, Nucl. Phys. **B 400** (1993) 161 [hep-th/9212149] [INSPIRE].

[25] D.J. Gross and W. Taylor, *Two-dimensional QCD is a string theory*, Nucl. Phys. **B 400** (1993) 181 [hep-th/9301068] [INSPIRE].

[26] D.J. Gross and W. Taylor, *Twists and Wilson loops in the string theory of two-dimensional QCD*, Nucl. Phys. **B 403** (1993) 395 [hep-th/9303046] [INSPIRE].

[27] N. Caporaso, L. Griguolo, M. Mariño, S. Pasquetti and D. Seminara, *Phase transitions, double-scaling limit and topological strings*, Phys. Rev. **D 75** (2007) 046004 [hep-th/0606120] [INSPIRE].

[28] N. Caporaso, M. Cirafici, L. Griguolo, S. Pasquetti, D. Seminara and R.J. Szabo, *Topological strings and large N phase transitions. II. Chiral expansion of q-deformed Yang-Mills theory*, JHEP **01** (2006) 036 [hep-th/0511043] [INSPIRE].

[29] B. Forbes and M. Jinzenji, *Local mirror symmetry of curves: Yukawa couplings and genus* 1, Adv. Theor. Math. Phys. **11** (2007) 175 [math.AG/0609016] [INSPIRE].

[30] E. Witten, *Two-dimensional gauge theories revisited*, J. Geom. Phys. **9** (1992) 303 [hep-th/9204083] [INSPIRE].

[31] E. Witten, *On quantum gauge theories in two-dimensions*, Commun. Math. Phys. **141** (1991) 153 [INSPIRE].

[32] M. Mariño, *Les Houches lectures on matrix models and topological strings*, CERN-PH-TH-2004-199, (2004) [hep-th/0410165] [INSPIRE].

[33] P.V. Buividovich, G.V. Dunne and S.N. Valgushev, *Complex path integrals and saddles in two-dimensional gauge theory*, Phys. Rev. Lett. **116** (2016) 132001 [arXiv:1512.09021] [INSPIRE].

[34] S. Giombi and V. Pestun, *The* 1/2 *BPS 't Hooft loops in $N = 4$ SYM as instantons in* 2*d Yang-Mills*, J. Phys. **A 46** (2013) 095402 [arXiv:0909.4272] [INSPIRE].

[35] B. Fiol and G. Torrents, *Exact results for Wilson loops in arbitrary representations*, JHEP **01** (2014) 020 [arXiv:1311.2058] [INSPIRE].

[36] E. Witten, *Theta dependence in the large N limit of four-dimensional gauge theories*, Phys. Rev. Lett. **81** (1998) 2862 [hep-th/9807109] [INSPIRE].

[37] *Digital Library of Mathematical Functions (DLMF) eq.* (18.15.22), http://dlmf.nist.gov/18.15.

[38] *Digital Library of Mathematical Functions (DLMF) eq.* (18.27.15), http://dlmf.nist.gov/18.27.

[39] R. Dijkgraaf, R. Gopakumar, H. Ooguri and C. Vafa, *Baby universes in string theory*, Phys. Rev. **D 73** (2006) 066002 [hep-th/0504221] [INSPIRE].

[40] M. Aganagic, H. Ooguri and T. Okuda, *Quantum entanglement of baby universes*, Nucl. Phys. **B 778** (2007) 36 [hep-th/0612067] [INSPIRE].

[41] K. Okuyama and K. Sakai, work in progress.

[42] M. Aganagic and K. Schaeffer, *Refined black hole ensembles and topological strings*, JHEP **01** (2013) 060 [arXiv:1210.1865] [INSPIRE].

[43] Z. Kökényesi, A. Sinkovics and R.J. Szabo, *Refined Chern-Simons theory and (q,t)-deformed Yang-Mills theory: semi-classical expansion and planar limit*, JHEP **10** (2013) 067 [arXiv:1306.1707] [INSPIRE].

JHEP04(2018)059